

上位キャッシュサイズを考慮した二重キャッシュ環境における下位キャッシュの置換手法 A Server Cache Size Aware Cache Replacement Algorithm for L2 Cache

長廻 雄介[†] 山口 実靖[†]
Yusuke Nagasako Saneyasu Yamaguchi

1. はじめに

近年、ストレージ容量は増大を続けており、扱うデータ量も増加の一途を辿っている。それに伴う管理コストの増加が情報社会の問題の一つとなっている。そこでNAS (Network Attached Storage) や SAN (Storage Area Network) といったネットワークストレージを用いたデータの集約・一元管理により、管理コストを削減する手法が提案され広く使われている。

ネットワークストレージへのアクセスは、サーバ計算機上のキャッシュとストレージ機器のキャッシュを介して行われる。このとき、サーバ側のキャッシュ置換手法がLRUである場合、ネットワークストレージに到着するI/O要求には通常とは逆向きの負の参照の時間的局所性が存在し、従来の参照の時間的局所性を期待するLRUキャッシュ置換手法は効果的に機能しないことが既存研究[1]から分かっている。

本稿では、この負の参照の時間的局所性を考慮したキャッシュ置換手法を提案する。

2. 負の参照の時間的局所性

ネットワークストレージへのアクセスは、図1のようにサーバ計算機キャッシュとストレージ機器キャッシュを介して行われる。サーバ計算機のキャッシュ置換手法は多くの場合LRU置換手法が用いられており、この場合最近参照されたデータはサーバ計算機のキャッシュに格納される。このため、最近アクセスされたデータへのアクセスはサーバ計算機上で処理され、ストレージ機器に要求が届くことはない。従って、ネットワークストレージへの参照要求には最近アクセスされたデータが近い将来再度参照されることが無い、という負の参照の時間的局所性が存在し、従来の参照の時間的局所性を期待しているLRUキャッシュ置換手法は効果的に機能しない。

図2はシミュレーションにより計測した再参照間隔と参照確率の関係性を示したものである。

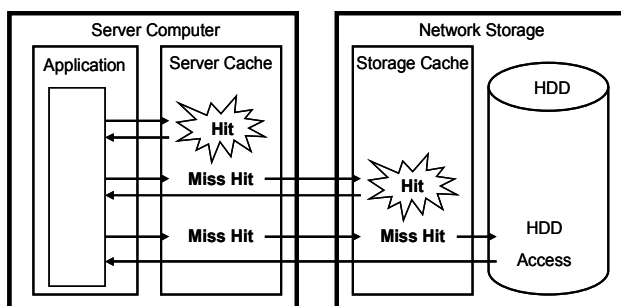


図1 二重キャッシュ構造

[†] 工学院大学大学院工学研究科電気・電子工学専攻
Electrical Engineering and Electronics, Kogakuin University
Graduate School

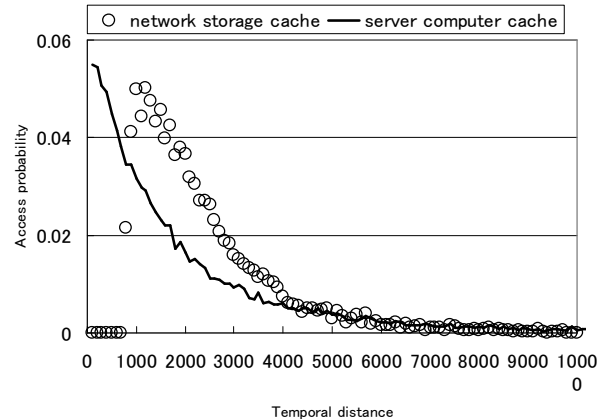


図2 負の参照の時間的局所性

横軸は一度参照されたデータが再度参照されるまでの時間(単位はアクセス数)を表し、縦軸はその時間でのアクセス確率を表している。サーバ計算機上へのアクセスは再度参照されるまでの時間が短いほどアクセス確率が高く、従来の時間的局所性が見られるが、ストレージ機器への再参照要求は近い将来(約1000アクセス以下)には発生しないことが分かる。これはシミュレーション上で設定したサーバ計算機のキャッシュサイズが、約1000アクセスで満たされるためであり、サーバ計算機のキャッシュと同等のアクセスがこない限り一度格納されたデータはサーバ計算機のキャッシュに残り、同一データへの再アクセスはサーバ計算機で処理されるためである。以上からネットワークストレージ上には一度参照されたデータは近い将来再度参照される可能性が低いという負の参照の時間的局所性が存在していることが分かる。

3. 提案手法

負の参照の時間的局所性を考慮した手法として、INTEを提案する。INTEではアクセス履歴を保持し、最も価値の無いデータをこの履歴から算出し選択する。前章で示したように、参照してからの経過時間が短い場合データはサーバ計算機キャッシュに存在するため、ストレージ機器を参照することなく、サーバキャッシュからデータが破棄され始めるとアクセス確率が上昇していく。このようにネットワークストレージへのアクセスは参照してからの経過時間によりアクセス確率が変動する。そこで提案手法では、アクセス履歴から図2のような再参照間隔と参照確率のPDFを作成し各データの価値を決定する。

各データの価値は、履歴中の最新アクセスからの経過時間から、ストレージキャッシュに残っているであろう時間までのPDFを積分した値とする。このため各データの価値とは、今後そのデータの予測ヒット確率であるといえる。

本稿では、ストレージキャッシュに残っていると予想される時間とはストレージキャッシュが満たされるアクセス数と設定している。しかし、ストレージキャッシュサイズがサーバキャッシュサイズ以下の場合、負の参照の時間的局所性によりアクセス直後のデータの積分範囲内はアクセス確率がすべて0になってしまうため、効果的に機能しない。よって、積分範囲は最低でもサーバキャッシュサイズの2倍とした。この条件のためサーバキャッシュサイズは既知である必要がある。

4. 評価

4.1 評価手法

提案手法の有効性を確認するため、シミュレーションによる評価を行った。本シミュレーションではアプリケーションプログラムからランダムにデータアクセス要求を発生させ、それぞれのキャッシュ置換アルゴリズム使用時のストレージキャッシュヒット率を調査した。ストレージサイズ(データ総数)は、1,000,000ブロック、サーバキャッシュサイズは1,024ブロック、ストレージ機器キャッシュサイズは512~4096ブロック、アクセス回数は500,000アクセス、アクセス対象データは無作為に選択した点をピークとする平均100kと256kの指数分布の偏りを持つ乱数により決定した。ただし、50,000アクセスごとにピークを変更する。

評価に用いたキャッシュ置換アルゴリズムは(1)LRU, (2)MQ, (3)FIX, (4)WC, (5)提案手法の5種類である。(1)LRUは最後のアクセスからの時間が最長のものを置換対象とする手法、(2)MQ(Multi Queue)[2]は2次キャッシュ(例えばストレージ機器のキャッシュ)を考慮した置換手法で、複数のLRUリストを管理し、頻度ごとに各リストに割り当て最下位リストの最も古いデータを破棄する手法、(3)は無作為に選択したデータを保持し続ける手法、(4)WC(Weighted Count)[3][4]は過去に我々が負の参照の時間的局所性を考慮した手法として提案した手法である。提案手法同様、再参照間隔をもとに各データの価値を決定する手法で、アクセスしてからの経過時間が0からサーバキャッシュサイズ分までは価値を1とし、経過時間がサーバキャッシュサイズと同等の時に価値を10として算出する。以降は指数関数に基づき価値を減少させる。各データの価値は上記したアクセスごとの価値の合計であり、その値が最も低いデータを破棄対象とする。WCは負の参照の局所性を考慮しており再アクセス直後のデータの価値を低く見積もっているが、サーバ計算機キャッシュから破棄された後のPDFの形状は考慮していない。(5)は提案手法である。

4.2 評価結果

図3, 図4はストレージキャッシュのヒット率を示している。どちらのアクセス分布においても提案手法が最もヒット率が高いことが確認できる。またLRUはストレージ機器キャッシュサイズが1000以下の場合キャッシュヒット率はほぼ0であり、ストレージ機器キャッシュサイズと比較してサーバキャッシュサイズが大きくなる限り、LRUは効果的に機能していないことが分かる。

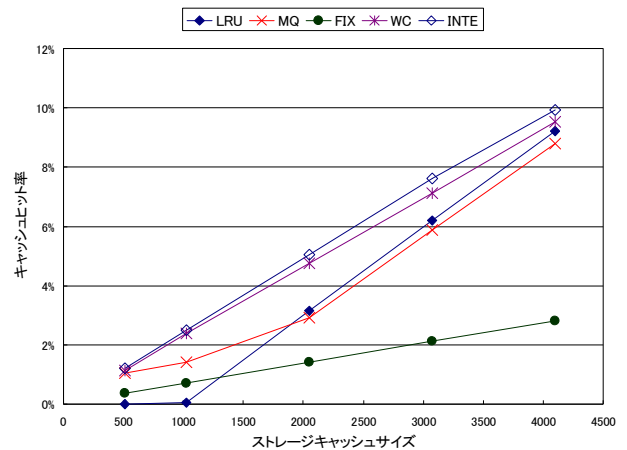


図3 ストレージキャッシュヒット率(平均 100k)

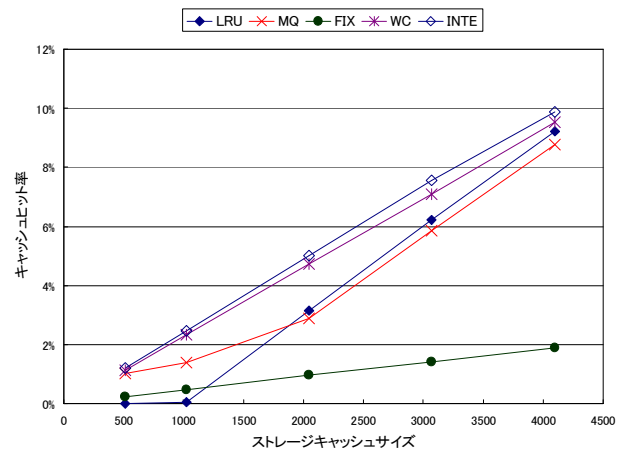


図4 ストレージキャッシュヒット率(平均 256k)

5. おわりに

本稿では負の参照の時間的局所性を考慮した手法として、INTE手法を提案し、シミュレーションによりその有効性を示した。今後は、書き込み処理に関する考察、キャッシュ置換手法のOSへの実装などを行う予定である。

謝辞

本研究は科研費(22700039)の助成を受けたものである

参考文献

- [1]宮野 晋平, 山口 実靖, 浅谷 耕一, “多段キャッシュ型ネットワークストレージへのアクセスの時間的局所性を考慮したメモリキャッシュ制御”, 情報処理学会研究報告 マルチメディア通信と分散処理研究会報告, 2009年3月
- [2]Yuanyuan Zhou and James F. Philbin, Kai Li “Second-Level Buffer Cache Management,” IEEE Transactions on parallel and distributed systems, vol. 15, no. 7, JULY 2004
- [3]長廻 雄介, 山口 実靖, “上位キャッシュの動作を考慮した下位キャッシュの置換手法”, 第73回(平成23年)情報処理学会講演論文集, 第73回分冊1 pp179.
- [4]Yusuke Nagasako and Saneyasu Yamaguchi, “A Server Cache Size Aware Cache Replacement Algorithm for Block Level Network Storage,” In Proceedings of The 4th Ad Hoc, Sensor and P2P Networks Workshop (AHSP) 2011